ArcSDE Geodatabase Replication
1-Way Replicas merged to a central location

Environmental Systems Research Institute's (ESRI) documentation on geodatabase replication explains common scenarios where you would establish one or more Arc Spatial Database Engine (SDE) geodatabase replicas to distribute data to other ArcSDE servers in the organization. The scenario of a "one-way" replication to distribute data from a database where records are edited to other sites (optionally with a geographic filter for that site's area of interest) is well covered. However, the reverse data flow is certainly possible as well, especially when data are managed at several sites – often with custom workflows or slightly varied schema requirements – and then are rolled up to a read-only database at the central site. ESRI has since published a technical article on establishing such a replication. *See* ESRI article #34200 *available at* http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=34200).

Basic Concepts

- The child replica must be established atSDE.Default version. The parent does not have to be created from the SDE.Default version.
- Allowing editors from non-SDE States requires that the child replica at the National Operations Center (NOC) be versioned to allow editing. According to ESRI, you must therefore use the "full replication" option instead of "simple replication."
- According to ESRI, "You must be connected as the owner of the data on the target Geodatabase when creating the replica." For this reason, it is safe to assume that the synchronization process is going to be initiated from the central location as it would have the data owner's account information. The central site must then have edit access to the parent/State office data; otherwise, the replication cannot be established. If the State site uses a protected document, there will be no way for the central site (NOC) to manipulate the data. Only by creating a version like any other editor and posting that data, can the partent/State ArcSDE data be manipulated.
- To achieve a full compress of the database, one must synchronize *all* replicas!

ESRI issues a word of caution for one-way replicas:
"If implementing a one way system where you sometimes need to edit the child replica's data, consider using two way replication. Since one-way replication assumes that the data are read only on the child, a synchronization may over-write edits to the child replica's data. The conflict detection logic of two way replication will flag these differences as conflicts allowing you to decide how to handle the differences. Two way replication allows data exchange in both directions, but also works in cases where you only send changes in one direction."

Creating Replicas

Following the ESRI documentation, all of the data should be assembled in the central/child replica ArcSDE instance. If you are working with topology and boundaries that are shared, it is strongly recommended that edits to correct topology errors and boundary discrepancies be completed before establishing the replication. Each state/parent replica will be a subset of the full dataset and not necessarily have the neighboring features for reference. This will make edge-matching that much more difficult,and resulting topology errors or boundary slivers will only be revealed when data are synchronized and validated in the central/child replica's ArcSDE database.

Set-up the merged dataset: Because of the potential for schema variations and even projection differences at the different replica sites, consider the following:

- Build a complete, merged dataset at the central/child ArcSDE site.
- The data should be in the projection in which it will be stored.
- If there are additional fields used by a given State/parent site, these need to be removed prior to building the replica. If there is a linking field present, it may be used to re-attach the data at the State site once the data are returned for loading immediately prior to building the replica. Alternatively, field deletions could be coordinated among all sites once it is loaded.
- After loading to the central/child ArcSDE instance, create global IDs on all replicated feature classes and tables. ArcCatalog has a context menu to do this, or you can use the tool in ArcToolbox -> Data Management Tools -> Distributed Geodatabase. Apply the appropriate privileges for editors that will not be editing via a replica.

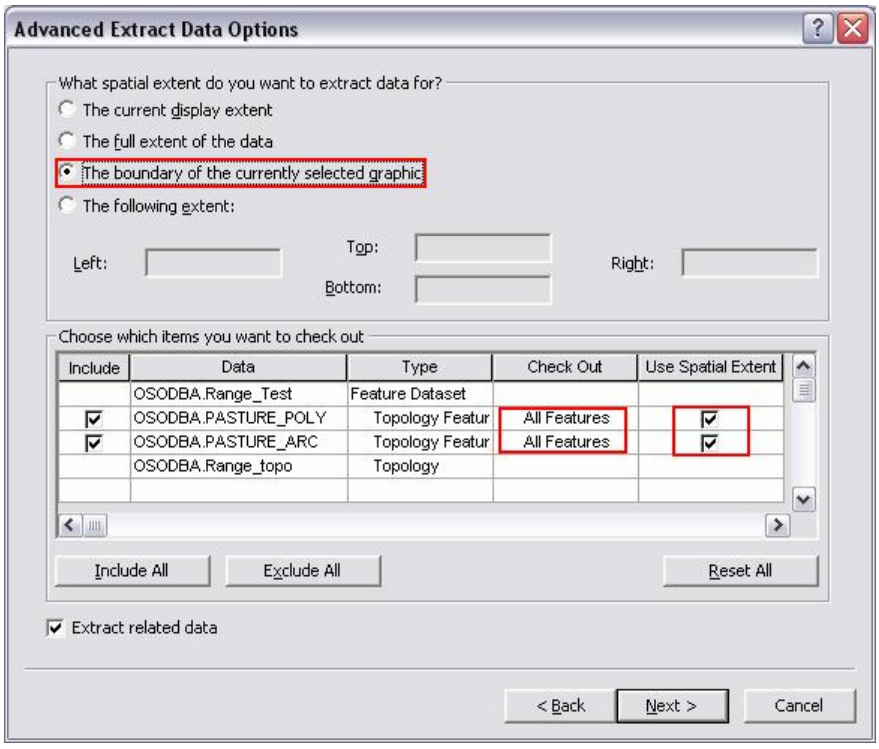
Create the replica datasets: For each State, use ArcMap to create a subset of the data for which they will be responsible. The ESRI documentation advises using the "Extract Data" tool from the Distributed Geodatabase toolbar. Note that there are some issues to take into consideration (at least at ArcGIS 9.2sp4):

- Do not attempt to make an extract of a selected set. Although this seems ideal (ex. select the polygons, then boundary arcs via shared segment option), the tool only outputs data from one of the feature classes that has a selection present. The other features are ignored but no error is returned and the resulting replica only has data in one feature class.
- Extracting directly to another geodatabase does not provide an option to select a configuration keyword. Unless you are using the DEFAULT keyword, extract the data to XML and load that to the State ArcSDE database and specify the keyword during the XML import dialog.
- Although not specifically addressed in ESRI's documentation, testing has shown that the parent and child replicas can be in different projections.

Start ArcMap, then add the data from the central/child ArcSDE instance that will be utilized in the replica. Create a graphic shape to delineate the extract area (this will extract all features that have a full or partial overlap with the shape). On the Distributed Geodatabase toolbar, start the Extract Data tool:

Check the "Data" radio button, and be sure to select the advanced options in the checkbox near the bottom. Extracting to an XML document is probably preferred.

Selecting the advanced options allows you to use the spatial extent specified by the graphic shape:



You should now have an extract of the features for your State ArcSDE geodatabase in a XML file to be transferred to the State site.

If your replica is in a different projection from the source data, you should now re-project the source data to the desired projection.

The State site should load the XML to the State database with the appropriate data owner account (specify the SDE configuration keywords if necessary).

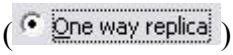
Each State should make any schema changes necessary. This may include deleting fields specific to other States, or temporary fields for linking.

- Add fields (and load with data prior to versioning to minimize system load if possible). In this case, a temporary spatial join feature class was created in the SDE database and used to update the feature class via SQL:

```
SQL> UPDATE pasture_poly f
  2  SET (f.altpast, f.grazsys) =
  3    (SELECT t.altpast_1, t.grazsys_1
  4     FROM tmp_range_join t
  5     WHERE t.objectid = f.objectid);

459 rows updated.

SQL> UPDATE pasture_poly f
  2  SET version_name = 'IntialLoad'
  3     WHERE version_name IS NULL;

459 rows updated.
```

- Dropping fields that are used in the central ArcSDE schema is not recommended as the values in the central database will remain populated, even after synchronization.

The State data owner will then register the data as versioned and apply privileges, including edit access to the account that the central/child site will use to build the replica. If there is a topology present, validate it (you may get some errors because the "extract by graphic" method pulls out all data that intersects the graphic polygon). If you have features in the State database that are outside of your area of responsibility due to the nature of the extract process, delete them and revalidate the topology. You should not be making any changes to the shapes or attributes at this point.

The central/child replica owner should create the one-way replica as explained in ESRI KB article #34200. (

):

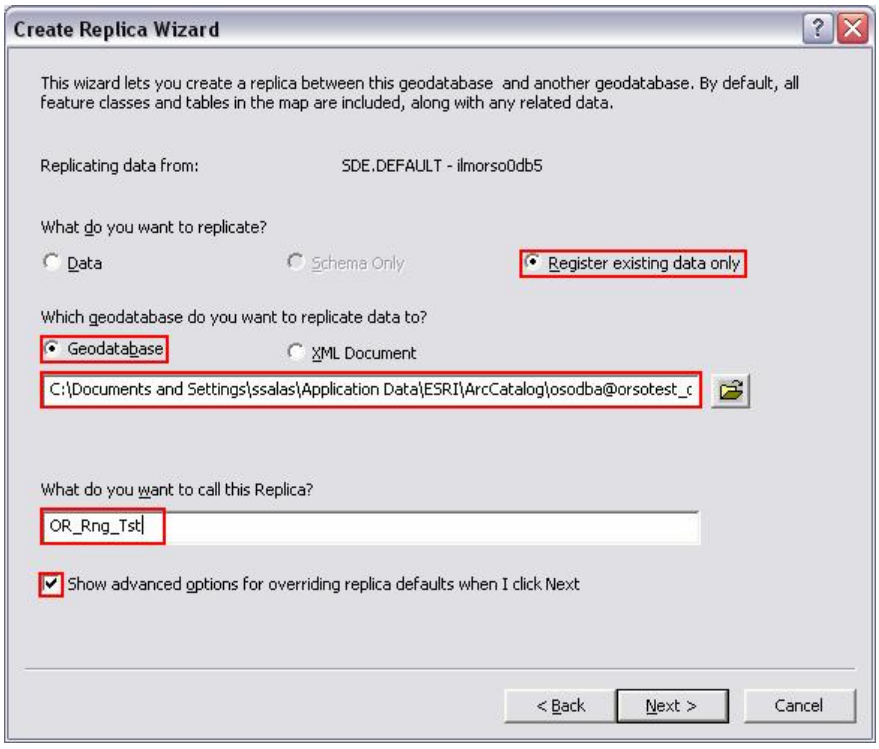
- Using a spatial filter is a good idea to prevent any features created at the State/parent replica sites that are outside of that site's area of responsibility from making it into the central ArcSDE database. This is a graphic shape created from the ArcMap Drawing toolbox and must be selected prior to opening the replica creation tool from the Distributed Geodatabase toolbar. Make the shape large enough to account for any data that will be subsequently added; otherwise, the replica will probably need to be recreated.

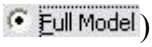
- Load the data from the State office ArcSDE into ArcMap as the user who will own the replica. This user must have edit rights to the data, even if he/she is not editing (replication utilizes versioning).

- Create the replica using the "Register existing data only" option. The SDE.Default version can be protected at this time, but keep in mind that synchronizing later will require that the Default version be temporarily made public during the synchronization process.

- Remember that 2-tier SDE connections "direct connect" generate more network load, so a 3-tier connection may be a better option.

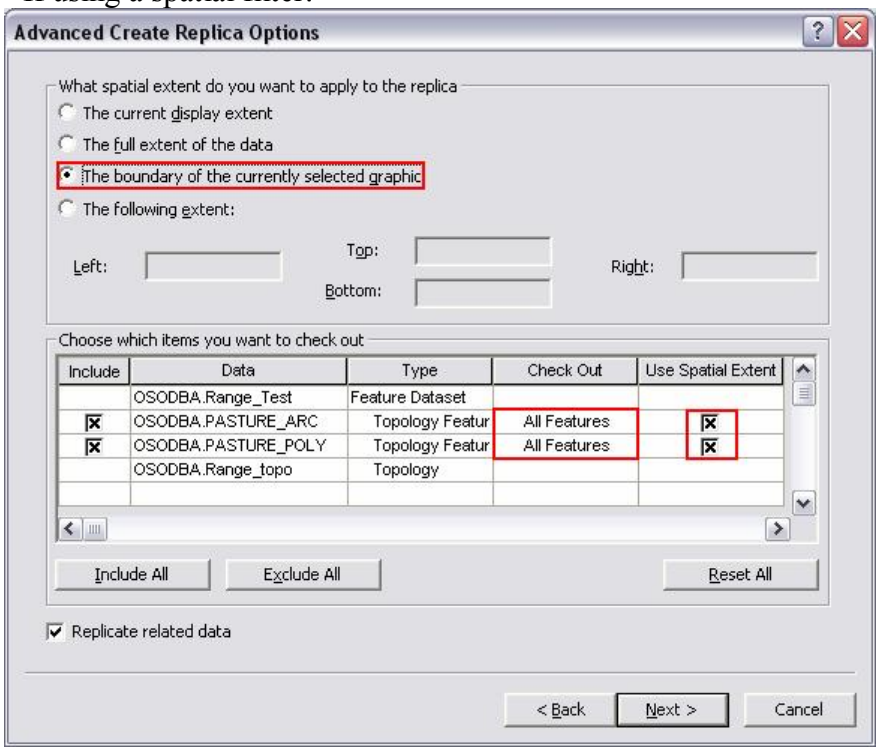
- Check the advanced options boxions



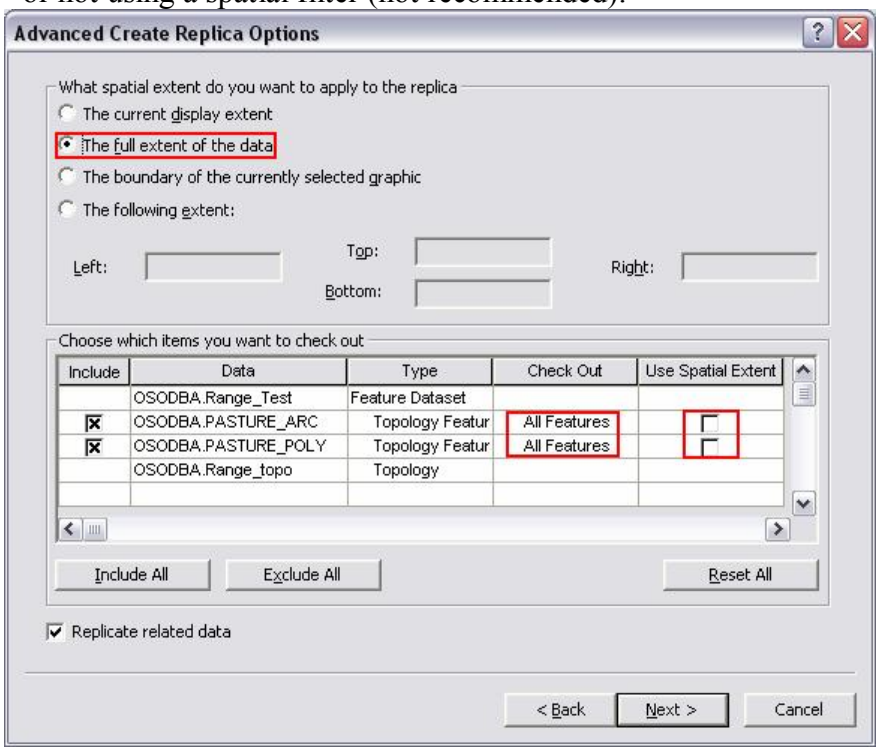
- Specify that you want a full replication (

). This permits versioned edits at the central site for those States that do not have ArcSDE replicas.

- If using a spatial filter:



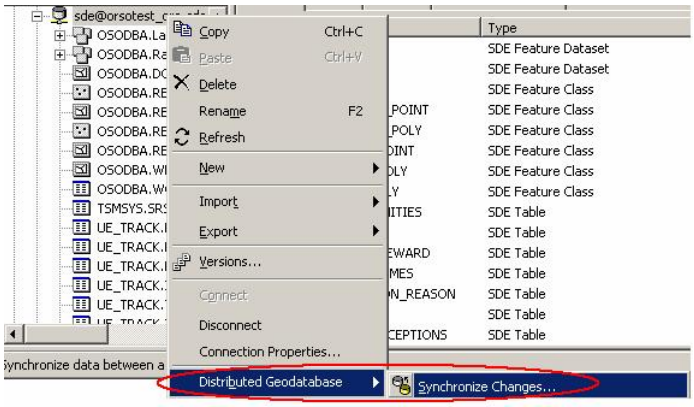
- or not using a spatial filter (not recommended):

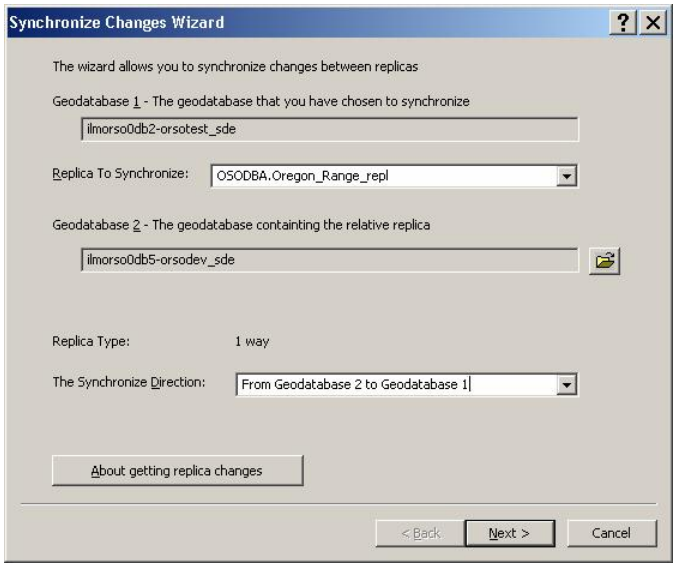


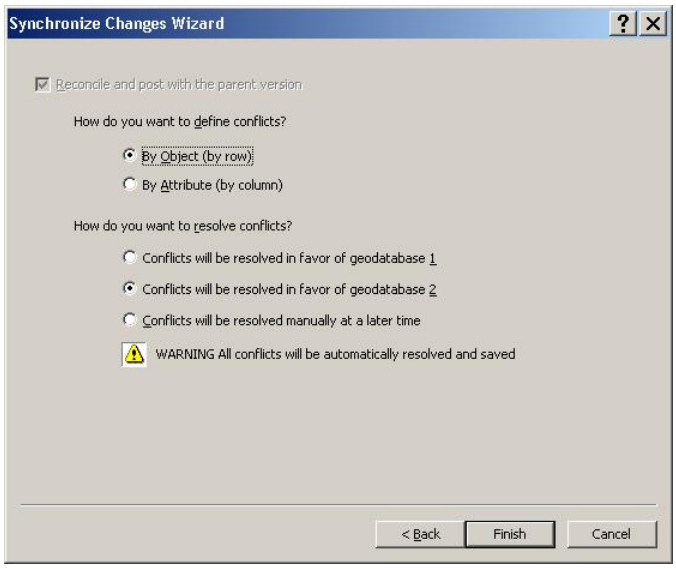
Using the Replicas

Once the replicas are established, each State site should be able to make edits utilizing their own workflows and processes. In the meantime, those sites without ArcSDE can connect to the central ArcSDE database and make changes directly (the central site's quality control should not permit changes to be made to data that are coming from a State/parent replica site; if it did, those changes would be lost later in favor of the replica State's changes). When the central site chooses to update its replica data, it will use the synchronization function to transfer changes from each State ArcSDE to its own. Keep in mind that this is important to do on a regular basis, even in the absence of changes, to allow for more complete database compressions.

The following screenshots show a manual synchronization in ArcCatalog. Another screenshot shows how the SDE.Default version at the central/child replica site can be changed to Public status – this is required for the replica synchronization to work.

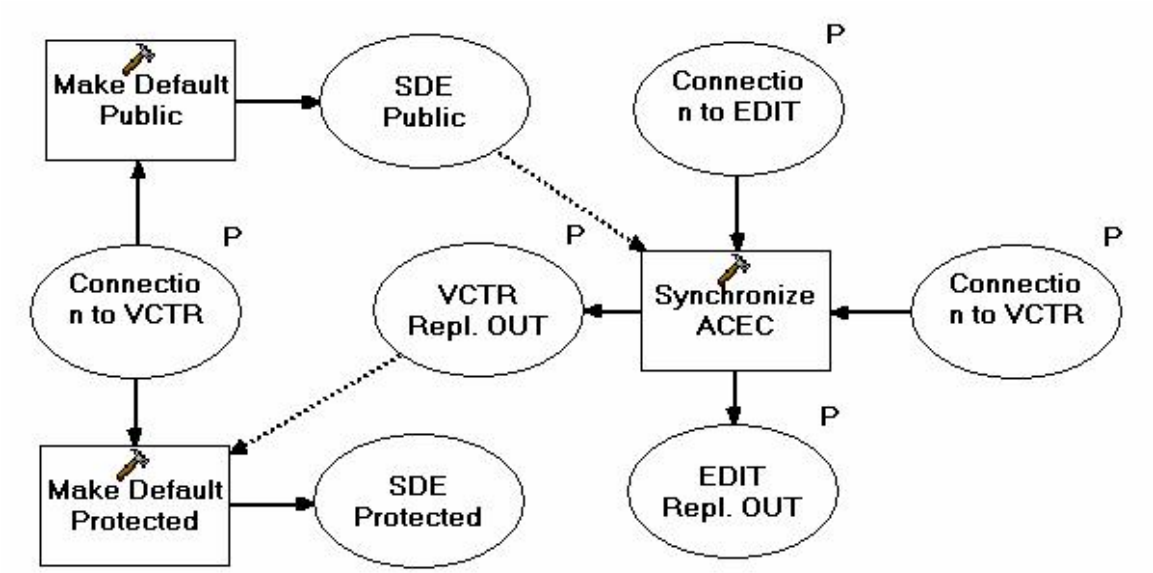

Copy Ctrl+C
Paste Ctrl+V
Delete
Rename F2
Refresh
New
Import
Export
Versions...
Connect
Disconnect
Connection Properties...
Distributed Geodatabase
Synchronize Changes...
Type
SDE Feature Dataset
SDE Feature Dataset
SDE Feature Class
SDE Feature Class
SDE Feature Class
SDE Feature Class
SDE Feature Class
SDE Table
SDE Table
SDE Table
SDE Table
SDE Table
SDE Table
SDE Table


Synchronize Changes Wizard
The wizard allows you to synchronize changes between replicas
Geodatabase 1 - The geodatabase that you have chosen to synchronize
ilmorso0db2-orsotest_sde
Replica To Synchronize:
OSODBA.Oregon_Range_repl
Geodatabase 2 - The geodatabase containting the relative replica
ilmorso0db5-orsodev_sde
Replica Type:
1 way
The Synchronize Direction:
From Geodatabase 2 to Geodatabase 1
About getting replica changes
< Back
Next >
Cancel


Synchronize Changes Wizard
Reconcile and post with the parent version
How do you want to define conflicts?
By Object (by row)
By Attribute (by column)
How do you want to resolve conflicts?
Conflicts will be resolved in favor of geodatabase 1
Conflicts will be resolved in favor of geodatabase 2
Conflicts will be resolved manually at a later time
WARNING All conflicts will be automatically resolved and saved
< Back
Finish
Cancel

The same operation can be performed with the ArcToolbox Geoprocessing tools. It is recommended that a model or script be built to make this easily repeatable and to even run as a scheduled task during non-peak-use hours. For example, the following model has 3 input parameters:

- the replica owner's ArcSDE connection at the State/parent site;
- the replica owner's ArcSDE connection at the central/child site; and
- the SDE user's ArcSDE connection at the central/child site.



The model uses the SDE user to switch the SDE.Default version from Protected to Public, runs the synchronization tool, and then returns SDE.Default to Protected status.

Changing the SDE.Default version at the central/child replica site to Public status